

# Limited preparation contextuality in quantum theory and its relation to the Cirel'son bound

Manik Banik, Some Sankar Bhattacharya, Amit Mukherjee, and Arup Roy

Physics and Applied Mathematics Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata 700108, India.

Andris Ambainis and Ashutosh Rai

Faculty of Computing, University of Latvia, Raina bulv. 19, Riga, LV-1586, Latvia.

Kochen-Specker (KS) theorem lies at the heart of the foundations of quantum mechanics. It establishes impossibility of explaining predictions of quantum theory by any noncontextual ontological model. Spekkens generalized the notion of KS contextuality in [Phys. Rev. A **71**, 052108 (2005)] for arbitrary experimental procedures (preparation, measurement, and transformation procedure). Interestingly, later on it was shown that preparation contextuality powers parity-oblivious multiplexing [Phys. Rev. Lett. **102**, 010401 (2009)], a two party information theoretic game. Thus, using resources of a given operational theory, the maximum success probability achievable in such a game suffices as a *bona-fide* measure of preparation contextuality for the underlying theory. In this work we show that preparation contextuality in quantum theory is more restricted compared to a general operational theory known as *box world*. Moreover, we find that this limitation of quantum theory implies the quantitative bound on quantum nonlocality as depicted by the Cirel'son bound.

Quantum mechanics (QM) departs fundamentally from the well known *local-realistic* world view of classical physics. This stark contrast of quantum theory with classical physics was illuminated by J. S. Bell [1]. Since the Bell's seminal work, nonlocality remains at the center of quantum foundational research [2, 3]. More recently quantum nonlocality has been also established as a key resource for device independent information technology [3, 4]. Quantum nonlocality does not contradict the relativistic causality principle, however, QM is not the only possible theory that exhibits nonlocality along with satisfying the no-signaling principle; there can be non-quantum no-signaling correlations exhibiting nonlocality. One extreme example of such a correlation (more non-local than QM) was first constructed by Popescu and Rohrlich (PR) [5]. Whereas PR correlation violates the Bell-Clauser-Horne-Shimony-Holt (Bell-CHSH) [1, 6] inequality by algebraic maximum, the optimal Bell-CHSH violation in quantum theory is restricted by the Cirel'son bound [7]. In this work, we show that Cirel'son limit on nonlocal behavior of quantum theory can be explained from its another very interesting feature, namely restricted *preparation contextuality*.

Nearly at the same time of Bell's result, Kochen and Specker proved another important no-go theorem showing that predictions of *sharp* (projective) measurements in QM can not be reproduced by any non-contextual ontological model [8]. Unlike Bell-nonlocality, structure of QM is implicit in the definition of KS contextuality. However, recently the idea of KS contextuality has been generalized, by Spekkens [9], for arbitrary operational theories rather than just quantum theory and for arbitrary experimental procedures rather than just sharp measurements. It was then shown that mixed preparations (density matrices) in quantum theory exhibit preparation contextuality [9, 12]. Interestingly, invoking

another non-classical concept called *steering* [10, 11] along with this new idea of preparation contextuality one can establish nonlocality of QM without using any Bell type inequalities; it has been shown that nonlocality of some hidden variable models, underlying QM, directly follows from the steerability of bipartite pure entangled states and the preparation contextuality of mixed states [12–14].

The traditional definition of contextuality address to only the contexts of projective measurements, which has been studied in much depth [15, 16]. However, generalized notion of contextuality developed by Spekkens define three different types of contexts: measurement (generalized), preparation, and transformation contexts [9]. Interest in studying contextuality in this general framework is relatively new and growing [12–14, 17–20]. This generalized approach has lead to designing more robust experimental tests of contextuality [18–20]; these recent results are very promising given that previously requirements for testing contextuality in experiments has been a topic of much controversy (for a more discussion see the ref. [20] and relevant references therein).

Our work here is build upon the notion of preparation contextuality which address the impossibility of representing two equivalent preparation procedures, of an operational theory, equivalently in some ontological model. More precisely, suppose we have two equivalent operational preparations  $P, P'$ , i.e., the outcome probabilities  $p(k|P, T, M) = p(k|P', T, M) \forall$  outcomes  $k$ , transformations  $T$ , and measurements  $M$ . Then a hidden variable (ontic) model, which reproduces  $p(k|P, T, M)$  by averaging over the ontic states  $\lambda$  is preparation non-contextual if  $\forall M, T : p(k|P, T, M) = p(k|P', T, M) \Rightarrow p(\lambda|P) = p(\lambda|P')$ , where  $p(\lambda|P)$  and  $p(\lambda|P')$  represent respective distributions over the ontic states followed by operational preparations  $P$  and  $P'$  [9].

Preparation contextuality has operational usefulness as it powers *parity-oblivious multiplexing* (POM), a two-party secure computation task [17]; in this work, Spekkens *et al.* derived a ‘noncontextuality inequality’ which place an upper bound on any operational theory that admits a preparation noncontextual ontological model. Further the authors showed that the success rate of the POM game played with only classical resources is restricted by the same inequality. Thus, Spekkens and coauthors concluded that any operational theory is preparation contextual if it can beat the classical bound in a POM task.

It turns out that, in performing a POM task, certain quantum resources can do better than any classical resource, thus proving that QM is a preparation contextual theory. Interestingly, in this work we show that for performing a POM task, there exists an operational (toy) theory, namely *box world* [21, 22], which can do better than QM. Thus, though QM is preparation contextual, the amount of preparation contextuality in QM is constrained compared to the box world. Furthermore, we show that restricted preparation contextuality of quantum theory leads to its limited nonlocal behavior as depicted in the Cirel’son bound. Therefore, our result brings the qualitative connection between preparation contextuality and nonlocality explored in [12–14] to a quantitative footing.

*Parity-oblivious multiplexing:* It is a variant of the well studied information-processing task called *random access code* (RAC) [23–25]. Suppose a  $n$ -bit string  $x$ , chosen uniformly at random from  $\{0, 1\}^n$ , is given to Alice. An integer  $y$ , chosen uniformly at random from  $\{1, 2, \dots, n\}$ , is given to Bob, now task for Bob is to guess the  $y^{\text{th}}$  bit of Alice’s input. Let us denote Bob’s guess as  $\beta_y$ . In the POM game Alice and Bob collaborate to optimize the guessing probability  $p(\beta_y = y^{\text{th}} \text{ bit of Alice})$ . Alice can send to Bob any information which encodes her input. However, there is a cryptographic constraint: no information about any parity of  $x$  can be transmitted to Bob. More specifically, letting  $s \in \text{Par}$  where  $\text{Par} \equiv \{r | r \in \{0, 1\}^n, \sum_i r_i \geq 2\}$  is the set of  $n$ -bit strings with at least 2 bits that are 1, no information about  $x \cdot s = \bigoplus_i x_i s_i$  (termed the  $s$ -parity) for any such  $s$  can be transmitted to Bob (here  $\oplus$  denotes sum modulo 2).

The main result of Spekkens *et al.* [17] can be now stated more precisely: *for  $n$ -bit POM game played with states (resources) from a preparation non-contextual theory, the average success probability is bounded as follows:*

$$p_{\text{NC}}(\beta_y = y^{\text{th}} \text{ bit of Alice}) \leq \frac{1}{2} \left(1 + \frac{1}{n}\right).$$

Motivated by this result, in our work, we define the maximum success probability in a POM task in an operational theory as a *bona-fide* measure to quantify the strength of preparation contextuality of the concerned

theory. The approach we adopt here is similar to defining the strength of nonlocality of correlations as the amount of Bell-CHSH violation (or maximum success probability in a Bell-CHSH game). In remaining of this paper, we focus on 2-bit POM task. For the 2-bit POM game we adopt two different schemes: (1) encoding-decoding scheme and (2) correlation assisted scheme.

(1) *Encoding-decoding scheme:* Alice and Bob can perform POM task by using resources of an general operational theory. In an operational theory, the primitives of description are preparations and measurements (for simplicity, here we do not consider dynamics/transformation of the system) [21, 22, 26–30]. The theory simply provides an algorithm for calculating the probability  $p(k|P, M)$  of an outcome  $k$  of measurement  $M$  given a preparation (state)  $P$ . The collection of all states, that the system can be prepared in, forms a compact and convex subset  $\Omega$  of a finite dimensional vector space  $V$ . Results of a measurement on any state  $\omega$  of the theory is described by an *effect*  $e : \Omega \rightarrow [0, 1]$ , which is a map such that  $e(\omega)$  is the probability of obtaining the outcome  $e$ . There is an *unit effect*  $u$  such that  $u(\omega) = 1 \forall \omega \in \Omega$ . Any measurement can now be expressed as some set of effects  $\{e_i\}$  such that  $\sum_i e_i = u$ .

Alice depending on the input string  $x \in \{00, 01, 10, 11\}$ , given to her uniformly at random, implements a preparation procedure  $P_x \in \Omega_A$  in an operational theory  $\mathcal{T}$  and sends the encoded particle to Bob. For each integer  $y \in \{1, 2\}$ , Bob implements a binary-outcome measurement  $M_y$ , and reports the outcome as his output. The *average* probability of winning is given by:

$$p_{\mathcal{T}}(\beta_y = y^{\text{th}} \text{ bit of Alice}) \equiv p(\beta_y = x_y) = \frac{1}{8} \sum_{y=1}^2 \sum_{x \in \{00, 01, 10, 11\}} p(\beta_y = x_y | P_x, M_y). \quad (1)$$

The optimal success probability in an operational theory is  $p_{\mathcal{T}}^{\text{opt}}(\beta_y = x_y) := \max_{P_x, M_y} p(\beta_y = x_y)$ , where optimization is performed over all possible encodings and decodings procedure allowed in the theory  $\mathcal{T}$ . Of course the encoding and decoding must satisfy the parity-oblivious constraint expressed here as:

$$p(P_{00}|k, M) + p(P_{11}|k, M) = p(P_{01}|k, M) + p(P_{10}|k, M); \quad \forall M \in \mathcal{M}, \text{ and } \forall k. \quad (2)$$

First we show that optimal success probability in box world is strictly greater than that of quantum theory, i.e.,  $p_{\text{box}}^{\text{opt}}(\beta_y = x_y) > p_Q^{\text{opt}}(\beta_y = x_y)$ . To prove this result we first consider the quantum case, and then the box world.

(a) *Quantum theory:* Alice encodes her 2 bits into the four pure qubits with Bloch vectors  $\{(\pm 1, 0, 0), (0, 0, \pm 1)\}$  equally distributed on the equatorial  $X - Z$  plane of the Bloch sphere; as shown in

Fig.(2). Bob performs the measurement  $(\sigma_x + \sigma_z)/\sqrt{2}$  if he wishes to learn the first bit, and the measurement  $(\sigma_x - \sigma_z)/\sqrt{2}$  if he wishes to learn the second. He guesses the bit value 0 upon obtaining the positive outcome, otherwise he guesses the bit value 1. In all cases, the guessed value is correct with probability  $\frac{1}{2}(1 + \frac{1}{\sqrt{2}})$ , which results the average success probability  $p_Q(\beta_y = x_y) = \frac{1}{2}(1 + \frac{1}{\sqrt{2}}) > \frac{2}{3} = p_{NC}^{opt}(\beta_y = x_y)$ . Since the parity 0 and parity 1 mixtures in this protocol are represented by the same density operator, no information about the parity can be obtained by any quantum measurement. Interestingly, the qubit protocol just described turns out to be quantum optimal.

**Proposition-1:** In a 2-bit POM game, optimum average success probability over all quantum encoding-decoding schemes is:  $p_Q^{opt}(\beta_y = x_y) = \frac{1}{2}(1 + \frac{1}{\sqrt{2}})$ .

*Proof:* Alice prepares and encodes as  $\{ij \mapsto \rho_{ij} : i, j \in \{0, 1\}\}$ , where  $\rho_{ij}$  are state operators acting on  $\mathbb{C}^d$ ; she can always find an appropriate pure state  $|\Psi_{12}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$  and projectors  $P_A, P_{A'}$  such that:  $\frac{1}{2}\rho_{00} = \text{tr}_1\{(P_A \otimes I)|\Psi_{12}\rangle\}$ ,  $\frac{1}{2}\rho_{11} = \text{tr}_1\{((I - P_A) \otimes I)|\Psi_{12}\rangle\}$ ,  $\frac{1}{2}\rho_{01} = \text{tr}_1\{(P_{A'} \otimes I)|\Psi_{12}\rangle\}$ ,  $\frac{1}{2}\rho_{10} = \text{tr}_1\{((I - P_{A'}) \otimes I)|\Psi_{12}\rangle\}$ . Alice performs one of the following projective measurements: (i)  $P_A \otimes I + (I - P_A) \otimes I = I \otimes I$ , (ii)  $P_{A'} \otimes I + (I - P_{A'}) \otimes I = I \otimes I$ , on part-1 of  $|\Psi_{12}\rangle$  and depending on the measurement result she sends part-2 to Bob (or discard it). Measurement (i) and (ii) respectively produces two decompositions  $\frac{1}{2}\rho_{00} + \frac{1}{2}\rho_{11}$  and  $\frac{1}{2}\rho_{01} + \frac{1}{2}\rho_{10}$  for part-2 of  $|\Psi_{12}\rangle$ . Alice in this way prepares and send  $\rho_{ij}$  to Bob; assured that parity obliviousness condition is always satisfied.

Bob on receiving part-2, performs a two outcome projective measurement  $\{P_B, (I - P_B)\}$  ( $\{P_{B'}, (I - P_{B'})\}$ ), if he is asked to guess Alice's first (second) bit, and answers 0(1) when measurement outcome is +1(-1). Here  $P_A, P_{A'}, P_B, P_{B'}$  are projectors acting on  $\mathbb{C}^d$ . Due to no constraint on the dimension of the Hilbert space, Neumark's theorem allows us to consider only projective measurements, without loss of generality. Substituting  $\rho_{ij}$  in terms of  $|\Psi_{12}\rangle$ ,  $P_A$  and  $P_{A'}$  in the expression for average success probability for the 2-bit POM game we get:  $p_Q = \frac{1}{8}[4 + \langle \Psi_{12} | \{A \otimes B + A' \otimes B + A \otimes B' - A' \otimes B'\} | \Psi_{12} \rangle]$ , where  $A = 2P_A - I$ ,  $A' = 2P_{A'} - I$ ,  $B = 2P_B - I$ ,  $B' = 2P_{B'} - I$  (see [31]-(A) for details). Since all four operators (observables)  $\{A, A', B, B'\}$  have eigenvalues  $\{\pm 1\} \in [-1, 1]$ , and any operator from the set  $\{A, A'\}$  commute with any operator from the set  $\{B, B'\}$ , by applying Cirel'son's result [7] it follows that  $\langle \Psi_{12} | \{A \otimes B + A' \otimes B + A \otimes B' - A' \otimes B'\} | \Psi_{12} \rangle \leq 2\sqrt{2}$ . This gives,  $p_Q \leq \frac{1}{2}[1 + \frac{1}{\sqrt{2}}]$ . We have already discussed that there exists quantum protocol to achieve this upper bound. ■

(b) *Box world:* Interestingly, one can exceed the optimal

quantum bound in the box world. This system can be understood as a black box taking a binary input  $x = 0, 1$  and returning a binary output  $a = 0, 1$  [22]. The state

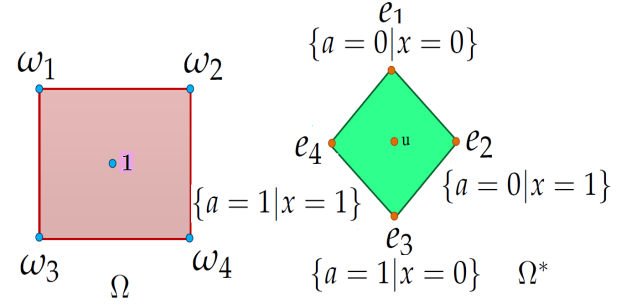


Figure 1. (Color on-line)  $\Omega$  denotes the normalized state space for the Box world. Four corners denote four deterministic states and the central dot denotes the completely mixed state.  $\Omega^*$  denotes the space of effects for the box world.  $\{e_i | i = 1, \dots, 4\}$  are four extremal effects of the Box world. Two ideal measurements are  $e_1 + e_3 = \mathbf{u} = e_2 + e_4$ .

of the system is described by a conditional probability distribution  $P(a|x)$ . The normalized state space  $\Omega$  of the system can be represented as a square in  $\mathbb{R}^2$  (see Fig.2). The system thus features four pure states  $\{\omega_j | j = 1, \dots, 4\}$ . For each pure state, the outcome 'a' is a deterministic function of the input 'x' ( $\omega_1 \rightarrow a = 0$ ,  $\omega_2 \rightarrow a = x$ ,  $\omega_3 \rightarrow a = 1$ , and  $\omega_4 \rightarrow a = x \oplus 1$ ). The center of  $\Omega$  is the maximally mixed state, that is, where 'a' is independent of 'x' and random. This maximally mixed state has non unique decomposition in term of pure states, i.e.,  $\frac{1}{2}\omega_1 + \frac{1}{2}\omega_3 = \frac{1}{2}\omega_2 + \frac{1}{2}\omega_4 = \mathbf{1}$ . The space of effects,  $\Omega^*$ , is the dual of  $\Omega$  (see Fig.2). It features four extremal effects  $\{e_j | j = 1, \dots, 4\}$  which correspond to the four measurement outcomes, i.e. obtaining output 'a' for a given input 'x'. The probability of  $e_j$  on any state is easily determined. For instance, effect  $e_1$  has probability one for states  $\omega_1, \omega_2$  and probability zero for states  $\omega_3, \omega_4$ . There are two pure measurements for this system: the first is composed of effects  $e_1$  and  $e_3$ , and corresponding to input  $x = 0$ ; the second is composed of effects  $e_2$  and  $e_4$ , and corresponding to input  $x = 1$ . Note that  $e_1 + e_3 = e_2 + e_4 = \mathbf{u}$ , where  $\mathbf{u}$  is the unit effect. This system is also known as generalized bit (g-bit) [21].

For performing the 2-bit POM with better than quantum success Alice and Bob pursue the following strategy in the box world. Alice encodes her strings as:

$$\begin{aligned} 00 &\rightarrow \omega_1, & 11 &\rightarrow \omega_3, \\ 01 &\rightarrow \omega_2, & 10 &\rightarrow \omega_4, \end{aligned} \quad (3)$$

and sends the encoded g-bit to Bob. The parity oblivious condition is satisfied as  $\frac{1}{2}\omega_1 + \frac{1}{2}\omega_3 = \frac{1}{2}\omega_2 + \frac{1}{2}\omega_4 = \mathbf{1}$ . To decode Alice's message, Bob performs (i) measurement  $\{e_1, e_3 | e_1 + e_3 = \mathbf{u}\}$  if he wishes to learn the first bit and (ii) measurement  $\{e_2, e_4 | e_2 + e_4 = \mathbf{u}\}$  if he wishes



to learn the second bit and he guesses Alice's bit as the measurement result. In every case, the guessed value is correct with certainty resulting  $p_{\text{box}}(\beta_y = x_y) = 1 > \frac{1}{2}(1 + \frac{1}{\sqrt{2}}) = p_Q^{\text{opt}}(\beta_y = x_y)$ . Clearly  $p_{\text{box}}^{\text{opt}}(\beta_y = x_y) = 1$  as the said strategy achieves 100% success probability.

(2) *Correlation assisted scheme*: Let Alice and Bob now follow a different scheme in which instead of sending encoded states they use correlations of pre-shared bipartite states allowed in an operational theory. In an operational theory a general bipartite correlation can be thought as a probability distribution  $p(\mathcal{C}, \mathcal{D}|\mathcal{U}, \mathcal{V}) \equiv \{p(c, d|u, v)\}$ , where  $u \in \mathcal{U}$ ,  $v \in \mathcal{V}$  are inputs given to Alice and Bob respectively and  $c \in \mathcal{C}$ ,  $d \in \mathcal{D}$  denote their respective outcomes. *No-signaling* correlations satisfy the conditions  $p(c|u) = \sum_{d \in \mathcal{D}} p(c, d|u, v)$ ;  $\forall c, u, v$  and the vice-versa. If local outcomes depend only on the choice of local measurements and (possibly) on the value of some shared (hidden) variable  $\lambda \in \Lambda$  which takes values according to some distribution  $p(\Lambda) = \{p(\lambda)\}$  then the correlation is called *local*, i.e.,  $p_L(c, d|u, v) = \sum_{\lambda \in \Lambda} p(\lambda)p(c|u, \lambda)p(d|v, \lambda)$  for all  $c, d, u, v$ . Correlations which can not be expressed in such form are called *nonlocal* [1]. Entangled quantum particles [32] exhibit nonlocal properties whereas they satisfy the no signaling conditions.

Using bipartite correlations Alice and Bob can perform the POM task in the following manner. Alice prior to start of the POM game, shares a correlated pair of particles, prepared in the state (preparation)  $P_{AB} \in \Omega_{AB}$ , with Bob. Depending on the input string, given to her, Alice performs measurement on her particle of the correlated pair and sends the measurement result to Bob via classical communications (CC). Bob, receiving CC from Alice, performs operations on his particle and tries to guess Alice bit. However the CC should not contain any information about the parity of the Alice's input string. It turns out that local correlations are not useful for performing the POM task.

**Proposition-2:** *Correlations having local description when assisted with classical communications are not useful for performing the parity-oblivious multiplexing task.*

*Proof:* Any local correlation between Alice and Bob can be thought as shared random variable  $\lambda \in \Lambda$  taking values according to a probability distribution  $p(\Lambda)$ . In Ref.[17], it has been shown that the only classical encodings of  $x$  that reveal no information about any parity (while encoding some information about  $x$ ) are those that encode only a single bit  $x_i$  for some  $i$ . For simplicity, without loss of generality, consider that the shared variable takes discrete values  $\lambda_i$  with  $\sum_{i=1}^n p(\lambda_i) = 1$ . If the variable takes value  $\lambda_k$  then Alice encodes her  $k^{\text{th}}$  bit and sends it to Bob. Bob, if asked, can correctly reveals this  $k^{\text{th}}$  bit while for him other bits ( $i \neq k$ ) are completely random. Thus using local correlation (shared randomness) Alice and Bob can design a strategy for de-

termining only one bit with certainty. But, whichever bit Bob guesses correctly, the (average) success probability of Bob's guess is bounded by the noncontextual bound  $\frac{1}{2}(1 + \frac{1}{n})$ . Therefore, due to convexity of distribution  $p(\Lambda)$ , it follows that any local correlation can not beat the noncontextual bound. ■

*Remark:* Here it is important to note that in a POM game to obtain greater success than the classical bound, the theory need not contain nonlocal correlations. For example, consider a theory in which individual state space is identical as quantum state space but the composite state space is severely restricted than quantum state space. The state space of the composite system is *minimal tensor product* [27] of individual Hilbert space and hence contains only separable states and hence no nonlocal correlation. In such a theory one can obtain the success probability of POM game as much as quantum theory by following the optimal encoding-decoding scheme of quantum theory. What the Proposition-2 proves is that if one wants to play the POM game by using correlation of such local theory she/he will not get any advantage.

**Proposition-3:** *Any no-signaling correlation  $\{p(ab|xy) : a, b, x, y \in \{0, 1\}\}$  violating Bell-CHSH inequality can exceed the classical bound for performing the 2-bit parity-oblivious multiplexing task. Moreover, if using a correlation average success in 2-bit POM game exceeds the quantum limit, then nonlocality of such correlation must exceed the Cirel'son bound.*

*Outline of Proof:* A proof follows from : (i) using the protocol for 2-bit random access code scenario discussed by Pawlowaski *et al.* [36] in the context of information causality, and (ii) showing that this protocol respect the parity obliviousness condition. Then this implies that, any nonlocal correlation can achieve more than the classical limit for 2-bit POM task. Moreover, the quantum limit for 2-bit POM game nonlocality of correlation is restricted by the Cirel'son bound. We give a complete proof of the proposition in the supplementary [31]-(C). ■

In quantum world using correlations of entangled particles Alice and Bob can win the POM game with more than classical (noncontextual) success probability. Using the steerability [10, 11] of the entangled state and the classical communication Alice tries to prepare Bob's state in different preparations depending on the input string given to her. For achieving the best result Alice attempts to prepare Bob's particle into states which achieve optimal success probability in the encoding-decoding scheme. Sharing two-qubit maximally entangled state Alice can prepare the optimal states by echoing an identical procedure as in the remote state preparation protocol [34] (see [31]-(B) for the protocol).

However, presence of steering, alone, in a theory is not sufficient for achieving more than classical success probability; the theory must also be preparation contex-

tual. For instance, there exists hypothetical *toy bit* theory [35] which allows steering, but the success probability of POM in this theory is restricted to the classical bound as the theory is preparation noncontextual (see [31]-(D)). On the other hand, though steerability in quantum theory is maximal, the optimal success probability of POM task is restricted due to its limited preparation contextuality.

To conclude, in this work we have considered an operational way to quantify the preparation contextuality of a general theory. We, then, show that quantum theory turns out to be less preparation contextual than another operational theory, namely box world. Further, we have shown that, in the quantum world, the restricted Bell-CHSH violation follows from limited preparation contextuality of the theory. Many researchers have tried to explain the limits of nonlocal feature in QM starting from a number of physically motivated ideas or principles. In particular, by considering various approaches, it has been successfully explained why Bell-CHSH quantity in quantum theory is restricted to Cirel'son bound [36]. Having established a link between the concept of nonlocality and preparation contextuality it would be interesting to suggest physical principal(s) leading to quantum bound on preparation contextuality.

Recently, the authors of [37] have shown that even-

parity-oblivious encodings are equivalent to the INDEX game, which implies  $2 \rightarrow 1$  POM game is equivalent to the well known Bell-CHSH nonlocal game. Therefore, a quantum encoding of  $2 \rightarrow 1$  POM with average success probability  $p_Q$  exists only if a quantum strategy for playing the Bell-CHSH game with the same average success probability exists. We take a different approach, by maximizing over all possible encoding-decoding schemes allowed in QM, we find the optimal success probability of 2-bit POM game; it turns out to be restricted compared to a more general operational theory. We conclude that restricted preparation contextuality, therefore, bounds the winning probability of Bell-CHSH game (nonlocality) in quantum theory.

*Acknowledgments:* Thanks to Guruprasad Kar for many stimulating discussions. MB like to acknowledge useful conversations with Rajjak Gazi. AM thanks Council of Scientific and Industrial Research, India for financial support through Senior Research Fellowship (Grant No. 09/093(0148)/2012-EMR-I). AsR likes to thank Andreas Winter for helping to prove the quantum optimality of 2-bit POM games. AA and AsR acknowledge support by the European Union Seventh Framework Programme (FP7/2007-2013) under the RAQUEL (Grant Agreement No. 323970) project, QALGO (Grant Agreement No. 600700) project, and the ERC Advanced Grant MQC.

- 
- [1] J. S. Bell, *Physics* **1**, 195–200 (1964).
  - [2] N. D. Mermin, *Rev. Mod. Phys.* **65**, 803 (1993).
  - [3] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, *Rev. Mod. Phys.* **86**, 419 (2014).
  - [4] V. Scarani, *Acta Physica Slovaca* **62**, 347 (2012) ([arXiv:1303.3081](#)).
  - [5] S. Popescu, D. Rohrlich, *Foun. Phys.* **24**, 379 (1994).
  - [6] J.F. Clauser, M.A. Horne, A. Shimony, R.A. Holt, *Phys. Rev. Lett.* **23**, 880 (1969).
  - [7] B. S. Cirel'son, *Lett. Math. Phys.* **4**, 93 (1980).
  - [8] S. Kochen and E.P. Specker, *Journal of Mathematics and Mechanics* **17**, 59 (1967).
  - [9] R. W. Spekkens, *Phys. Rev. A* **71**, 052108 (2005).
  - [10] E. Schrödinger, *Mathematical Proceedings of the Cambridge Philosophical Society* **32**, 446 (1936).
  - [11] H. M. Wiseman, S. J. Jones, and A. C. Doherty, *Phys. Rev. Lett.* **98**, 140402 (2007).
  - [12] M. Banik, S. S. Bhattacharya, S. K. Choudhary, A. Mukherjee, A. Roy, *Found. Phys.* **44**, 1230 (2014).
  - [13] N. Harrigan, R. W. Spekkens, *Found. Phys.* **40**, 125 (2010).
  - [14] M. S. Leifer, O. J. E. Maroney, *Phys. Rev. Lett.* **110**, 120401 (2013).
  - [15] A. Peres, *Phys. Lett. A* **151**, 107 (1990); N. D. Mermin, *Phys. Rev. Lett.* **65**, 3373 (1990); N. D. Mermin, *Rev. Mod. Phys.* **65**, 803 (1993).
  - [16] A. Grudka *et al.* *Phys. Rev. Lett.* **112**, 120401 (2014); S. Yu and C. H. Oh, *Phys. Rev. Lett.* **108**, 030402 (2012); A. Cabello, [arxiv:1201.0374](#).
  - [17] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, *Phys. Rev. Lett.* **102**, 010401 (2009).
  - [18] R. Kunjwal and R. W. Spekkens, [arxiv:1506.04150](#) (Accepted in *Phys. Rev. Lett.*).
  - [19] M. D. Mazurek *et al.* [arxiv:1505.06244](#).
  - [20] M. F. Pusey, [arxiv:1506.04178](#).
  - [21] J. Barrett, *Phys. Rev. A* **75**, 032304 (2007).
  - [22] P. Janotta, C. Gogolin, J. Barrett, and N. Brunner, *New J. Phys.* **13**, 063024 (2011).
  - [23] S. Wiesner, *SIGACT News* **15**, 78 (1983).
  - [24] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, [arXiv:quant-ph/9804043](#).
  - [25] A. Ambainis, A. Nayak, A. Ta-Shma, U. Vazirani, *Journal of the ACM*, **49**, 496 (2002).
  - [26] P. Busch, M. Grabowski, and P.J. Lahti, Springer-Verlag, Berlin, (1997).
  - [27] H. Barnum and A. Wilce, [arXiv:0908.2352](#).
  - [28] C. Pfister, [arXiv:1203.5622](#).
  - [29] N. Stevens, P. Busch, *Phys. Rev. A* **89**, 022123 (2014).
  - [30] I. Namioka, R.R. Phelps, *Pac. J. Math.* **31**, 469 (1969).
  - [31] See Supplemental Material at [URL will be inserted by publisher] for, (A) Proof of optimality of the 2-bit POM protocol, (B) An entanglement assisted scheme in QM, (C) Complete proof of the proposition-3, and (D) Example of a steerable toy bit theory not useful for POM games.
  - [32] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Rev. Mod. Phys.* **81**, 865 (2009).
  - [33] L. Masanes, A. Acín, and N. Gisin, *Phys. Rev. A* **73**, 012112

- (2006).  
 [34] C. H. Bennett, D. P. DiVincenzo, P. W. Shor, J. A. Smolin, B. M. Terhal, and W. K. Wootters, *Phys. Rev. Lett.* **87**, 077902 (2001).  
 [35] R. W. Spekkens, *Phys. Rev. A* **75**, 032110 (2007).  
 [36] M. Pawłowski *et al.* *Nature* **461**, 1101 (2009); M. Navascués, and H. Wunderlich, *Proc. Roy. Soc. Lond. A* **466**, 881 (2009); J. Oppenheim and S. Wehner, *Science* **330**, 1072 (2010); M. Banik *et al.* *Phys. Rev. A* **87**, 052125 (2013); A. Cabello, *Phys. Rev. A* **90**, 062125 (2014).  
 [37] A. Chailloux, I. Kerenidis, S. Kundu, J. Sikora, *arXiv:1404.5153*.

## SUPPLEMENTARY

### (A):- OPTIMALITY OF 2-BIT POM IN QUANTUM THEORY

Alice prepares and encodes as  $\{ij \mapsto \rho_{ij} : i, j \in \{0, 1\}\}$ , where  $\rho_{ij}$  are state operators acting on  $\mathbb{C}^d$ ; she can always find an appropriate pure state  $|\Psi_{12}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$  and projectors  $P_A, P_{A'}$  such that:

$$\begin{aligned}\frac{1}{2}\rho_{00} &= \text{tr}_1\{(P_A \otimes I)|\Psi_{12}\rangle\}, \\ \frac{1}{2}\rho_{11} &= \text{tr}_1\{((I - P_A) \otimes I)|\Psi_{12}\rangle\}, \\ \frac{1}{2}\rho_{01} &= \text{tr}_1\{(P_{A'} \otimes I)|\Psi_{12}\rangle\}, \\ \frac{1}{2}\rho_{10} &= \text{tr}_1\{((I - P_{A'}) \otimes I)|\Psi_{12}\rangle\}.\end{aligned}$$

Alice performs one of the following projective measurements:

- (i)  $P_A \otimes I + (I - P_A) \otimes I = I \otimes I$ , or
- (ii)  $P_{A'} \otimes I + (I - P_{A'}) \otimes I = I \otimes I$

on part-1 of  $|\Psi_{12}\rangle$  and depending on the measurement result she sends part-2 to Bob (or discard it). Measurement (i) and (ii) respectively produces two decompositions  $\frac{1}{2}\rho_{00} + \frac{1}{2}\rho_{11}$  and  $\frac{1}{2}\rho_{01} + \frac{1}{2}\rho_{10}$  for part-2 of  $|\Psi_{12}\rangle$ . Alice in this way prepares and send  $\rho_{ij}$  to Bob; assured that parity obliviousness condition is always satisfied.

Bob on receiving part-2, performs a two outcome projective measurements  $\{P_B, (I - P_B)\}$  ( $\{P_{B'}, (I - P_{B'})\}$ ), if he is asked to guess Alice's first (second) bit, and answers 0(1) when measurement outcome is +1(-1). Here  $P_A, P_{A'}, P_B, P_{B'}$  are projectors acting on  $\mathbb{C}^d$ . Due to no constraint on the dimension of the Hilbert space, Neumark's theorem allows us to consider only projective measurements, without loss of generality. The expression for average success probability for the 2-bit POM game thus become:

$$\begin{aligned}p_Q &= \frac{1}{8} [\text{Tr}\{\rho_{00}(P_B + P_{B'})\} + \text{Tr}\{\rho_{01}(P_B + (I - P_{B'}))\} + \text{Tr}\{\rho_{10}((I - P_B) + P_{B'})\} \\ &\quad + \text{Tr}\{\rho_{11}((I - P_B) + (I - P_{B'}))\}] \\ &= \frac{1}{8} [4 + \text{Tr}\{(\rho_{00} - \rho_{11})P_B\} + \text{Tr}\{(\rho_{01} - \rho_{10})P_B\} + \text{Tr}\{(\rho_{00} - \rho_{11})P_{B'}\} \\ &\quad - \text{Tr}\{(\rho_{01} - \rho_{10})P_{B'}\}], \\ &= \frac{1}{8} [4 + 2\{\langle\Psi_{12}|(2P_A - I) \otimes P_B|\Psi_{12}\rangle + \langle\Psi_{12}|(2P_{A'} - I) \otimes P_B|\Psi_{12}\rangle \\ &\quad + \langle\Psi_{12}|(2P_A - I) \otimes P_{B'}|\Psi_{12}\rangle - \langle\Psi_{12}|(2P_{A'} - I) \otimes P_{B'}|\Psi_{12}\rangle\}], \\ &= \frac{1}{8} [4 + \langle\Psi_{12}|\{(2P_A - I) \otimes (2P_B - I) + (2P_{A'} - I) \otimes (2P_B - I) \\ &\quad + (2P_A - I) \otimes (2P_{B'} - I) - (2P_{A'} - I) \otimes (2P_{B'} - I)\}|\Psi_{12}\rangle], \\ &= \frac{1}{8} [4 + \langle\Psi_{12}|\{A \otimes B + A' \otimes B + A \otimes B' - A' \otimes B'\}|\Psi_{12}\rangle],\end{aligned}$$

where  $A = 2P_A - I$ ,  $A' = 2P_{A'} - I$ ,  $B = 2P_B - I$ ,  $B' = 2P_{B'} - I$ .

### (B):- ENTANGLEMENT ASSISTED PROTOCOL IN QUANTUM MECHANICS

Using correlation of entangled quantum particles, assisted with classical communication, Alice and Bob can win the POM game with better than classical success probability while satisfying the parity oblivious condition.

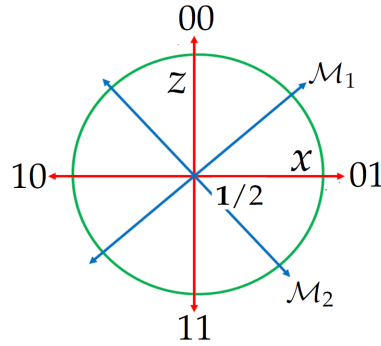


Figure 2. (Color on-line) The optimal quantum encoding-decoding scheme for playing the 2-bit POM. Green circle denotes the equatorial circle on  $X - Z$  plane of the Bloch sphere. Red and blue lines denote encoding and decoding scheme respectively.

For achieving the optimal success Alice, depending on her input string, will try to prepare Bob's particle in the states in which she had encoded her strings in the optimal *encoding-decoding* scheme (see Fig.2), i.e., she tries to prepares Bob's particles accordingly:

$$\begin{aligned} 00 &\rightarrow \frac{1}{2}(\mathbf{1} + \sigma_z), & 11 &\rightarrow \frac{1}{2}(\mathbf{1} - \sigma_z), \\ 01 &\rightarrow \frac{1}{2}(\mathbf{1} + \sigma_x), & 10 &\rightarrow \frac{1}{2}(\mathbf{1} - \sigma_x). \end{aligned}$$

Let Alice shares a correlated pair of particles with Bob prepared in singlet state  $|\psi^-\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |1\rangle_B - |1\rangle_A \otimes |0\rangle_B)$ . If the input sting is 00 or 11 then Alice performs  $\sigma_z$  measurement on her part of singlet pair otherwise she performs  $\sigma_x$  measurement. In both cases the *unconditional* state of Bob's particle is  $\frac{1}{2}\mathbf{1}$ . Hence by performing most general quantum measurement Bob can not get any information about the parity of the input string.

Whenever Alice's input string is 00 she performs  $\sigma_z$  measurement on her particle. If her measurement outcome is '1' then due to anti-correlation of singlet correlation Bob's particle is prepared in the state  $\frac{1}{2}(\mathbf{1} + \sigma_z)$ . But, if her measurement result is '0' then state of the Bob's particle is  $\frac{1}{2}(\mathbf{1} - \sigma_z)$  and it requires Bob to perform a rotation on his particle so that the state of the particle ends is the desired state  $\frac{1}{2}(\mathbf{1} + \sigma_z)$ . So, depending on her measurement result, Alice sends classical communication to Bob to inform whether he needs to perform the rotation or not. Note that this classical communication does not carry any information about the parity of the string. In rest of the cases Alice follows similar procedures and succeeds in preparing Bob's particle in the desired states. Note that this protocol echoes the same strategy as followed in the remote state preparation protocol [34].

For guessing Alice's bit, Bob performs the measurements as done in the optimal *encoding-decoding* scheme. For decoding the first bit he performs the spin measurement  $(\sigma_z + \sigma_x)/\sqrt{2}$  and for second bit he performs spin measurement  $(\sigma_z - \sigma_x)/\sqrt{2}$  and guesses the bit depending on the measurement result, which results in the  $\frac{1}{2}(1 + \frac{1}{\sqrt{2}})$  winning probability of the POM task.

### (C):- COMPLETE PROOF OF THE PROPOSITION-3 IN THE MAIN TEXT

**Proposition-3:** Any no-signaling correlation  $\{p(ab|xy) : a, b, x, y \in \{0, 1\}\}$  violating Bell-CHSH inequality can exceed the classical bound for performing the 2-bit parity-oblivious multiplexing task. Moreover, if using a correlation average success in 2-bit POM game exceeds the quantum limit, then nonlocality of such correlation must exceed the Cirel'son bound.

*Proof:* We start with a definition for nonlocality of correlation  $p(ab|xy)$ . The Bell-CHSH inequality expressed in terms of these correlations take a form:  $\mathbb{B} = \sum_{x=0}^1 \sum_{y=0}^1 p(a \oplus b = xy|x, y) \leq 3$ . Any correlation for which  $\mathbb{B} > 3$  is nonlocal, maximum nonlocality achievable in quantum mechanics is Cirel'son bound  $2 + \sqrt{2}$ . Over set of all

no-signalling correlation  $\mathbb{B}$  can achieve the maximum value 4. By a suitable local randomization technique used in [33], keeping nonlocality parameter  $\mathbb{B}$  invariant, one can write such no-signaling correlation in a canonical form:  $p(a \oplus b = xy|x, y) = \frac{1}{2}(1 + \gamma)$ ; where  $0 \leq \gamma \leq 1$ . For  $\gamma > 1/2$  correlations are nonlocal and  $\gamma = 1/\sqrt{2}$  correspond to Cirel'son bound  $\mathbb{B}_Q = 2 + \sqrt{2}$ .

Suppose, using canonical correlations, Alice and Bob play the 2-bit POM game with the same protocol as in 2-bit random access code scenario discussed by Pawłowski *et al.* [36] in the context of information causality. Let us denote Alice's two bit string as  $x_1 x_2$ , where  $x_1, x_2 \in \{0, 1\}$ . Alice inputs  $x = x_1 \oplus x_2$  and obtains  $a$  as outcome. Alice communicates  $c = x_1 \oplus a$  to Bob. If Bob is asked to guess Alice's first bit he gives  $y = 0$  as input otherwise he inputs  $y = 1$ . After obtaining outcome  $b$  from the shared canonical correlation and communication  $c$  from Alice he declares his answer as  $b \oplus c$ . With this protocol one can check that the average success probability of correct guess by Bob is  $\frac{1}{2}(1 + \gamma)$ . Note that the parity obliviousness condition is satisfied in this protocol due to: i) correlation used, being a no-signaling resource, can not carry any information, and ii) the communicated bit  $c$  contains no information about the parity of Alice's string. Now it is easy to see that, with this protocol, any nonlocal correlation achieves more than the classical limit for 2-bit POM task. Moreover, the quantum limit for 2-bit POM game implies that  $\gamma \leq 1/\sqrt{2}$  which can hold only if  $\mathbb{B}_Q \leq 2 + \sqrt{2}$ , i.e., nonlocality of correlation is restricted by the Cirel'son bound. ■

#### (D):- STEERABLE TOY-BIT THEORY IS NOT USEFUL FOR POM

Spekkens has introduced a toy theory in order to argue for an epistemic view of quantum states [35]. The theory is based on a principle, namely *knowledge balance principle*, according to which the number of questions about the physical state of a system that are answered must always be equal to the number that are unanswered in a state of maximal knowledge.

For the elementary system the number of questions in the canonical set is two, and consequently the number of *ontic* states is four. Denote the four ontic states as '1', '2', '3', and '4'. An *epistemic* state is nothing but a probability distribution  $\{(p_1, p_2, p_3, p_4) | p_i \geq 0 \forall i \text{ \& } \sum_{i=1}^4 p_i = 1\}$ , over the ontic states. Denoting disjunction by the symbol ' $\vee$ ' (read as *or*), the six possible pure *epistemic* states, allowed by the knowledge balance principle, read as:

$$\begin{aligned}
 1 \vee 2 &\leftrightarrow \left\{ \frac{1}{2}, \frac{1}{2}, 0, 0 \right\} = \begin{array}{|c|c|c|c|} \hline \blacksquare & \blacksquare & \square & \square \\ \hline \end{array}, \\
 3 \vee 4 &\leftrightarrow \left\{ 0, 0, \frac{1}{2}, \frac{1}{2} \right\} = \begin{array}{|c|c|c|c|} \hline \square & \square & \blacksquare & \blacksquare \\ \hline \end{array}, \\
 1 \vee 3 &\leftrightarrow \left\{ \frac{1}{2}, 0, \frac{1}{2}, 0 \right\} = \begin{array}{|c|c|c|c|} \hline \blacksquare & \square & \blacksquare & \square \\ \hline \end{array}, \\
 2 \vee 4 &\leftrightarrow \left\{ 0, \frac{1}{2}, 0, \frac{1}{2} \right\} = \begin{array}{|c|c|c|c|} \hline \square & \blacksquare & \square & \blacksquare \\ \hline \end{array}, \\
 1 \vee 4 &\leftrightarrow \left\{ \frac{1}{2}, 0, 0, \frac{1}{2} \right\} = \begin{array}{|c|c|c|c|} \hline \blacksquare & \square & \square & \blacksquare \\ \hline \end{array}, \\
 2 \vee 3 &\leftrightarrow \left\{ 0, \frac{1}{2}, \frac{1}{2}, 0 \right\} = \begin{array}{|c|c|c|c|} \hline \square & \blacksquare & \blacksquare & \square \\ \hline \end{array}.
 \end{aligned}$$

For such a system, one have less than maximal knowledge if both questions in the canonical set to be unanswered. This corresponds to the epistemic mixed state:

$$1 \vee 2 \vee 3 \vee 4 \leftrightarrow \left\{ \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4} \right\} = \begin{array}{|c|c|c|c|} \hline \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ \hline \end{array}.$$

The mixed state has following different convex decompositions:

$$1 \vee 2 \vee 3 \vee 4 = (1 \vee 2) +_{cx} (3 \vee 4), \quad (4)$$

$$= (1 \vee 3) +_{cx} (2 \vee 4), \quad (5)$$

$$= (1 \vee 4) +_{cx} (2 \vee 3). \quad (6)$$

The above set of decompositions can be thought as different preparation procedures for the same mixed state, a phenomenon that can be observed in quantum theory.

Compatible with knowledge balance principle, there are two types of (pure) epistemic states for a pair of elementary systems:



(1)  $(a \vee b).(c \vee d) \equiv (a.c) \vee (a.d) \vee (b.c) \vee (b.d)$ ; where  $a, b, c, d \in 1, 2, 3, 4$  and  $a \neq b, c \neq d$ .

(2)  $(a.e) \vee (b.f) \vee (c.g) \vee (d.h)$ ; where  $a, b, c, d, e, f, g, h \in \{1, 2, 3, 4\}$  and  $a, b, c, d$  are all different and same is for  $e, f, g, h$ .

For the second type of states the state for marginal elementary systems (both) is  $1 \vee 2 \vee 3 \vee 4$ . Let Alice shares with Bob a bipartite elementary system prepared in the state  $(1.1) \vee (2.2) \vee (3.3) \vee (4.4)$ . If Alice implements the measurement that distinguishes  $1 \vee 2$  from  $3 \vee 4$  on her part of the system then will be able to remotely prepare Bob system in decomposition of Eq.(4). Similarly implementing measurements that distinguishes  $1 \vee 3$  from  $2 \vee 4$  and  $1 \vee 4$  from  $2 \vee 3$  she can prepares other two decompositions, i.e., decomposition of Eq.(5) and Eq.(6) respectively, which establishes steering like phenomena for toy-bit theory.

A model is said to be preparation non-contextual if the probability distributions over the ontic states corresponding to different preparation procedures of an identical operational state remains identical. Otherwise the model will be called preparation contextual. It has been proved that any mixed quantum state is preparation contextual [9, 12]. But in this case, though the mixed state  $(1 \vee 2 \vee 3 \vee 4)$  can be prepared in different ways (decompositions (4), (5) and (6)) all the preparations give identical probability distribution  $\{\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\}$  over the ontic states  $\{1, 2, 3, 4\}$  for every preparation procedure. Therefore the theory is preparation non-contextual. Hence though the toy-bit theory exhibits steering phenomena, but being preparation noncontextual it gives no advantage in POM task.

---